Ensemble Classification of Hyperspectral Images using DWT and Support Vector Machine

K.Kavitha, G.Manochitra

Abstract: Remote sensing is the technique of acquiring information about the earth surface and identification of earth surface features to estimate the geographical properties using electromagnetic radiation as a medium of interaction. Remote sensing hyperspectral images provides the detailed analysis of the surface of the earth using advanced imaging instruments which can produce high-dimensional images with hundreds of spectral bands. Hyperspectral sensors collect information as a set of 'images'. In this paper, a unique technique is used for classifying the various classes present in the hyperspectral image and to produce a classification map. A transform called discrete wavelet transform is used in hyperspectral images as preprocessing. Statistical and co-occurrence features are used for extracting the features. Genetic Algorithm is used to solve optimization problems, capability of finding optimal solutions for high dimensional data. Genetic Algorithm gives best feature and the selected feature is used to initialize random populations. Selection, cross over and mutation is done to create new individuals. Support Vector Machine is used for classifying various classes. After ensemble the various classes, classification map is produced. Experiment is conducted with AVIRIS Indian Pines, Hyperion Botswana and ROSIS Pavia University. Results are validated with ground truth and accuracies are calculated and compared.

Keywords: Hyperspectral image, Discrete Wavelet Transform, Genetic algorithm, Support Vector Machine.

I. INTRODUCTION

Hyperspectral imagery deals with imaging narrow discrete spectral bands over spectral image and produce spectra of all pixels. It collects and process information from electromagnetic spectrum. Hyperspectral images with hundreds of bands offer high resolution and potential accuracy in land cover image classification. The process of relating pixels in an image to known class in an image is called Image Classification. The algorithm used for classification process is called Image Classifiers.

.

K.Kavitha is working as Assistant Professor in ECE Dept, Mepco Schlenk Engineering College, Sivakasi.
G.Manochitra is a PG student from Mepco Schlenk Engineering College, Sivakasi. Email:jaimanosankar@gmail.com

Discrete wavelet transform (DWT) gives better accuracy when used for image classification [1]. DWT is a time-scale approach that has been used successfully in many applications [2-3]. Ghazali implemented feature extraction technique using 2D-DWT and extracted co-efficient are used to represent image for classification [4].Chang explained that two-dimensional discrete wavelet transform has shown considerable promise in image denoising [5].

The high dimensionality of the hyperspectral data makes it more difficult to use it for classification. Moreover there is a lot of redundancy in the data which needs to be removed [6]. Complexity lies in the nature of high dimensionality data and the consequent ground truth demand for supervised classification [7].

In particular, this aspect known as Hughes phenomenon implies that the required number of labeled training samples for supervised classification increases as a function of dimensionality.

In remote sensing, the number of training samples available is often limited and this limitation becomes relevant in case of high number of features. This problem is addressed by identifying a model that is less sensitive to Hughes phenomenon provided it should reduce the redundancy of the dataset available.

The feature selection and extraction for SVM are also explained in detail by Battati [8]. The usage of SVM classifier for hyperspectral images is shown by Gualtieri [9]. Multiclass classifier for Hyperspectral Images is explained by Scholkopf [10].

In SVM classification, the accuracy is affected by whether the training data can provide a representative description of each class or not. In general, the more the number of pure training pixels, higher the classification accuracy can be obtained.

When classifiers have different accuracies, it is reasonable to give more discriminated power for those that have greater accuracy [11]. This is the idea for Majority Vote (MV) [12]. Such approaches are known as trainable because they need to find a set of weights to produce the best set of support for each sample. For combining the classifiers weighted linear combination genetic algorithm (WLC-GA) is used [13].

At first the hyperspectral image is transformed using discrete wavelet transform. Statistical features and co-occurrence features are used for feature extraction. Genetic algorithm is used for selecting the best features and the selected features are classified using support vector machine to produce classification map. Obtained classification map is compared with ground truth and accuracy is calculated for every class.

II. PROPOSED METHODOLOGY

A. Discrete Wavelet Transform

A Wavelet transform is the representation of a function by wavelets. The Wavelet transform uses wavelets of finite energy. Wavelet Transforms have advantages over traditional transforms for representing functions that have discontinuities and sharp peaks for accurate deconstructing and reconstructing images. Wavelet transform at high frequencies gives good time resolution and poor frequency resolution while at low frequencies gives good frequency resolution and poor time resolution.

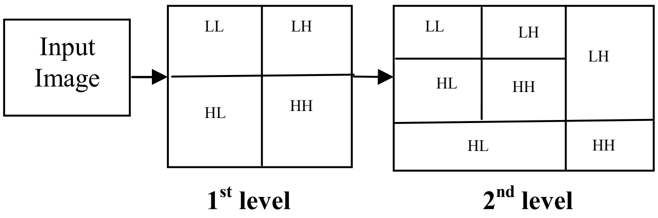


Fig 1: Concept of DWT

The two-dimensional DWT (2D-DWT) is a multi level decomposition technique and converts images from spatial domain to frequency domain. Applying DWT responds to processing of image by filters in each dimension for deconstructing. The filters divide the input image into four multi-resolution sub-bands LL, LH, HL, HH. The sub-band LL represents the approximate DWT coefficients while the sub-bands LH, HL and HH represent the horizontal, vertical and diagonal DWT coefficients. By applying inverse DWT, reconstruction of hyperspectral is obtained.

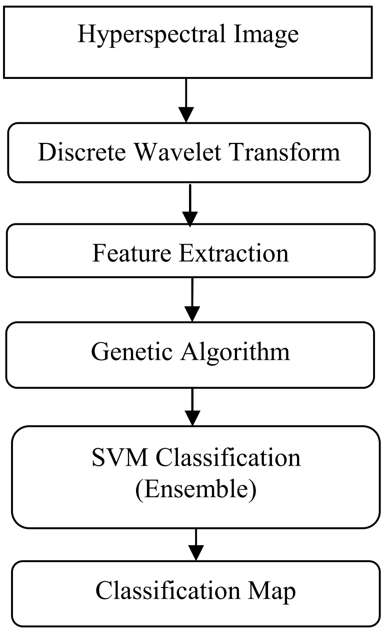


Fig 2: Proposed Work

B. Feature Extraction

After applying DWT to hyperspectral images, features are extracted using statistical and co-occurrence features. The formula for statistical and co-occurrence features are shown in table 1 and 2

Table 1 Statistical Features

Feature	Formula
Mean	$\mu_{ij} = (\sum_{1\le i\le M} \sum_{1\le j\le N} X_{ij}) / MN$
Variance	$\sigma_{ij} = (\sum_{1\le i\le M} \sum_{1\le j\le N} (X_{ij} - \mu_{ij})^2)^{0.5} / MN$
Standard Deviation	$\sigma_{ij}^2 = (\sum_{1\le i\le M} \sum_{1\le j\le N} (X_{ij} - \mu_{ij})^2) / MN$

Table 2 Co-occurrence Features

Feature	Formula
Energy	$\sum_i \sum_j P^2[i,j]$
Contrast	$\sum_i \sum_j (i-j)^2 P[i,j]$
Maximum Probability	{P[i,j]}
Cluster Tendency	$\sum_{i,j} (i+j-2\mu) kp[i,j]$

C. Genetic Algorithm

Genetic algorithms belong to the larger class of evolutionary algorithms (EA), which generate solutions to optimization problems using techniques inspired by natural evolution, such as mutation, selection, and crossover. The various steps involved in genetic algorithm are

1. Select the features for generating random population.
2. Evaluate the fitness function for the selected features
3. At first generation for every individual population random permutation is done.
4. To obtain next generation the cross-over probability should be greater than 0.5 to select best features.
5. Every individual values are calculated using their fitness function.
6. For mutation to create new individuals best values are selected to produce new generation.
7. After obtaining the best feature generation criteria is stopped.

D. SVM Classification

The classification is done with Support Vector Machines. A Support Vector Machine performs classification by constructing an N-dimensional hyper plane that optimally separates the data into two strategies. Given some training data D, a set of *n* points of the form

$$D = \{(x_i, y_i) \mid x_i \in R^p, y_i \in \{-1,1\}\}_{i=1}^{n} \quad (1)$$

Where the y_i is either 1 or −1, indicating the class to which the point X_i belongs. Each X_i is a *p*-dimensional real vector. To find the maximum-margin hyper plane that divides the points having $y_i = 1$ from those having $y_i = -1$. If the training data are linearly separable, select two hyper planes in a way that they separate the data and there are no points between them, and then try to maximize their distance. The region bounded by them is called "the margin". The hyper planes can be described by the equations 2 and 3

$$W.X - b = 1 \quad (2)$$

$$W.X - b = -1 \quad (3)$$

III. EXPERIMENTAL DESIGN

The experiment were conducted on AVIRIS Indian Pines, Hyperion Botswana and ROSIS Pavia University.

1. AVIRIS hyperspectral dataset is taken over the Indian Pines test site in North-western Indiana and consists of 145×145 pixels and 224 spectral reflectance bands in the wavelength range 0.4–2.5 10 μm. The Indian Pines scene contains two-thirds agriculture, one-third forest and other natural perennial vegetation. The ground truth available is designated into sixteen classes. The number of bands is reduced to 180 by removing bands covering the region of water absorption. The ground truth is also available for the dataset.
2. The Hyperion Botswana dataset is a sequence of data over the Okavango Delta, Botswana in 2001-2004. The Hyperion sensor EO-1 acquires data at 30 m pixel resolution over a 7.7 km strip in 145 bands covering the 400-2500 nm portion of the spectrum in 10 nm windows. The data consists of observations from 14 identified classes representing the land cover types in seasonal swamps, occasional swamps, and drier woodlands located in the distal portion of the Delta.
3. Pavia University scene was acquired by the ROSIS sensor during a flight campaign over Pavia University, northern Italy. The number of spectral bands is 103 for Pavia University. Pavia University has 610×340 pixels in each band. The geometric resolution is 1.3 meters. The ground truths differentiate 9 classes for Pavia University.

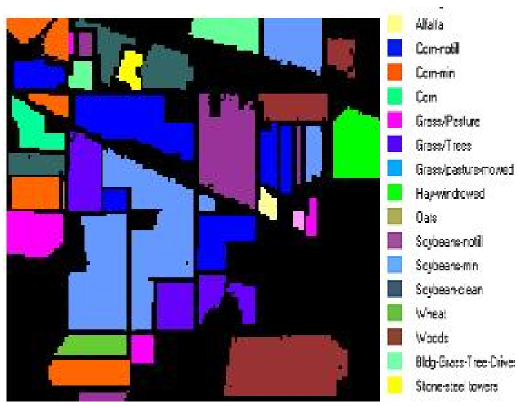
IV. RESULTS AND DISCUSSION

Hyperspectral image taken from AVIRIS, Hyperion, and ROSIS is chosen for the experiment. Hyperspectral image with lot of narrow bands offers high resolution. The dataset consists of 16 different classes for AVIRIS Indian Pines, 14 classes for Hyperion Botswana and 9 classes for Pavia University.

There are total 220 bands for Indian Pines dataset out of which 180 bands are chosen then for Botswana 50 bands are chosen out of 145 bands and for Pavia University 103 bands were chosen.

The classification accuracy depends on the feature extraction and classifier selection process. Given input image is transformed using Discrete Wavelet Transform. Discrete wavelet transform gives spatial and spectral information for the processed hyperspectral image. Discrete wavelet transform with first level and second level decomposition is performed for these datasets. Mean, variance, standard deviation, energy, contrast, maximum probability and cluster pixel tendency are used as statistical and co-occurrence features for discrete wavelet transform. Genetic Algorithm is applied to each feature. Genetic Algorithm selects the best feature and the feature is classified using Support Vector Machine to obtain classification map. Fitness function is calculated for each process to select best feature among the various features. For SVM classification pixels are randomly chosen from each class and their features are used for training. Here 5% of pixels from each class are used for training. All the pixels are tested against the training samples. The classifier produces the output, whether the pixel under test belongs to the interested trained class or not. Thus the pixels of interested classes are identified among the whole dataset. Similarly other classes are also trained.

(a) AVIRIS Indian Pines Input Image



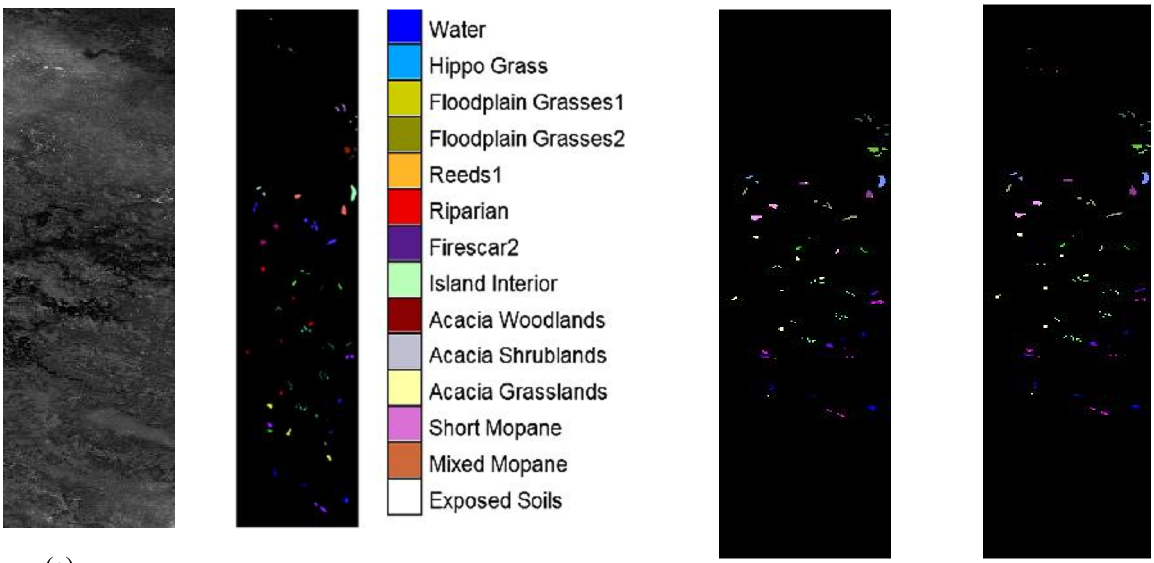
(b) Ground Truth Image

(c) 1st Level DWT

(d) 2nd Level DWT

Table 3 Accuracy Table for AVIRIS Indian Pines

	Class Name	Total Samples	Training Samples	Average Accuracy 1st level DWT (%)		Average Accuracy 2nd level DWT (%)	
				SVM	**GA-SVM**	**SVM**	**GA-SVM**
C1	Alfalfa	46	3	95.42	100	92.45	93.47
C2	Corn-Notill	1428	72	92.41	94.74	98.42	100
C3	Corn-Min	830	42	89.96	94.57	86.56	94.95
C4	Corn	237	12	68.47	93.24	92.31	94.69
C5	Grass-Pasture	483	24	92.13	95.03	87.42	93.67
C6	Grass-Trees	730	37	87.35	98.08	92.17	94.82
C7	Grass Pasture Mowed	28	2	75.63	82.14	92.36	98.63
C8	Hay-Windrowed	478	24	92.87	95.93	72.54	64.28
C9	Oats	20	1	94.68	95.39	98.56	99.37
C10	Soybean-Notill	972	49	92.49	98.97	85.46	89.28
C11	Soybean-Min	2455	123	96.31	98.78	92.57	99.27
C12	Soybean-clean	593	30	97.59	98.31	98.63	99.06
C13	Wheat	205	10	96.10	99.84	95.98	96.12
C14	Woods	1265	63	96.35	97.66	87.26	93.17
C15	Building Grass Trees Drives	386	19	92.22	94.62	95.69	100
C16	Stone Steel Tower	93	5	96.54	100	92.48	98.70
Over all Accuracy (%)				**91.03**	**96.08**	**91.30**	**97.85**

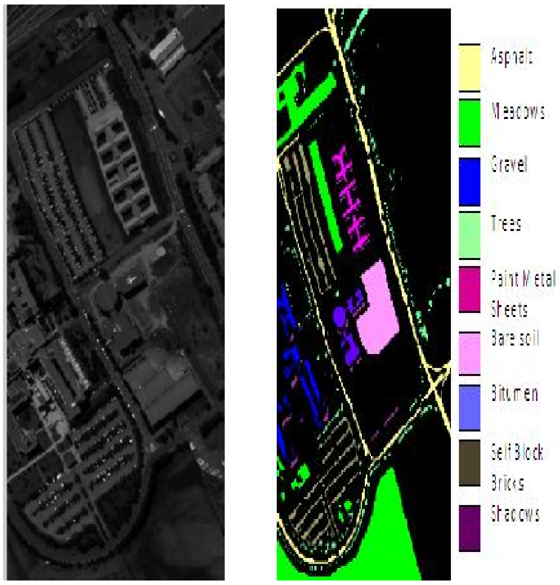


(a) (b)

Fig 4 (a) Hyperion Botswana Input Image (b)Ground Truth (c)1st Level DWT (d) 2nd Level DWT

Table 4 Classification output for Hyperion Botswana using 1st level DWT and 2nd level DWT

Class	**Class Name**	**Total Samples**	**Training Samples**	**Average Accuracy using 1st level DWT (%)**		**Average Accuracy using 2nd level DWT**	
				SVM	**GA-SVM**	**SVM**	**GA-SVM**
C1	Water	270	14	92.31	95.65	92.13	99.92
C2	Hippo Grass	101	5	97.52	100	94.23	95.16
C3	Flood Plain Grasses 1	251	13	91.30	95.02	94.75	95.30
C4	Flood Plain Grasses 2	215	11	86.54	94.12	92.63	93.24
C5	Reeds	269	14	83.21	99.06	92.96	95.23
C6	Riparian	269	14	91.85	93.67	96.58	97.94
C7	Fire scar	259	13	74.36	92.76	99.85	100
C8	Island Interior	203	10	90.12	92.19	92.54	95.18
C9	Acacia Woodlands	314	16	95.47	99.48	75.45	80.00
C10	Acacia Shrublonds	248	12	92.34	98.96	92.63	99.48
C11	Acacia Grasslands	305	15	92.58	93.67	96.32	98.85
C12	Short Mopane	181	9	91.57	92.41	96.74	97.80
C13	Mixed Mopane	268	13	75.02	85.46	92.48	97.92
C14	Exposed Soils	95	5	74.36	74.13	97.10	95.69
Overall Accuracy (%)				**87.75**	**93.32**	**93.31**	**95.83**

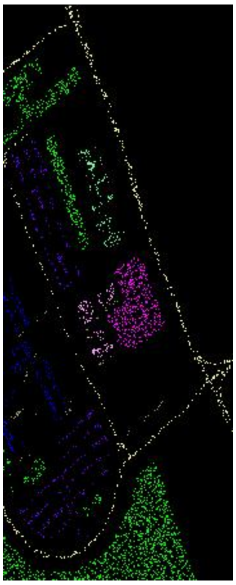


(a) (b)

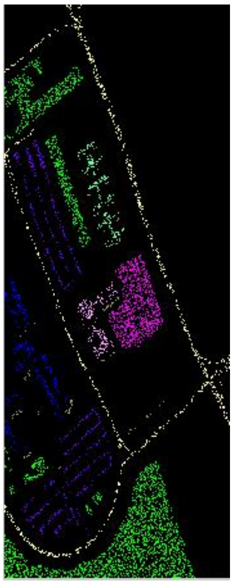
Fig 5 (a) ROSIS Pavia University Input Image (b) Ground Truth (c) 1st level DWT (d) 2nd level DWT

Table 5 Classification Output for ROSIS Pavia University using 1st level and 2nd level DWT

Class	Class Name	Total Samples	Training Samples	Average Accuracy for 1st level DWT		Average Accuracy for 2nd level DWT	
				SVM	GA-SVM	SVM	GA-SVM
C1	Asphalt	6631	331	76.52	91.05	81.24	91.85
C2	Meadows	18649	932	83.45	95.78	89.56	97.54
C3	Gravel	2099	105	67.20	92.02	75.85	85.78
C4	Trees	3064	153	78.58	93.79	75.87	93.99
C5	Metal sheets	1345	67	88.96	98.07	93.54	97.45
C6	Bare soil	5029	251	87.23	90.00	86.43	92.12
C7	Bitumen	1330	66	87.41	84.65	89.21	85.74
C8	Self blocking bricks	3682	184	82.45	89.25	82.78	90.52
C9	Shadow	947	47	65.52	80.54	76.20	87.68
Over all Accuracy (%)				79.48	90.68	83.40	91.40



(c)



(d)

V. CONCLUSION

In this paper a unique technique is used for classifying hyperspectral images. Discrete wavelet transform is used for transforming the hyperspectral images. Statistical and co-occurrence features are used for extracting the features. Genetic algorithm is used for selecting the features and support vector is used for classifying the various classes and by combining the different classes classification map is obtained. By using GA-SVM the overall accuracy using 1st level DWT for Indian Pines is 96.08% and by using 2nd level DWT overall accuracy us 97.85%. For Hyperion Botswana using 1st level DWT it produces overall accuracy as 93.32% and 95.83% using 2nd level DWT. For ROSIS Pavia university using 1st level DWT the overall accuracy is 90.68% and using 2nd level DWT overall accuracy is 91.40%.

REFERENCES

[1] Kamarul Hawari Ghazali, Mohd. Marzuki Mustafa, Aini Hussain, "Image Classification using Two dimensional Discrete Wavelet Transform", International conference on Instrumentation, control & automation,pp.71-74,2009

[2] A.popov, A. kanaykin, M.Zhukov, O. Panichev, O.Bodilovsky, "Adapted Mother Wavelets for Identification of Epileptiform complexes in Electroencepahalograms" Electronics and Electrical engineering.-Kaunas:Technologija, no.8-pp.89-92,2010

[3] V.M. Georgieve "An approach for computed Tomography Images Enhancement" Electronics and Electrical Engineering. –Kaunas:Technologija, no.2.pp 71-74, 2008

[4] Ghazali K.H, Mansor M.F, Mustafa m.M "Feature Extraction Technique using Discrete Wavelet Transform for Image Classification" Research and development,2007 SCOReD 2007. 5th student Conference on December 2007

[5] S.G.Chang, Bin Yu, M.Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," IEEE Trans. Image Processing, Vol.9,pp.1522-1531, sept.2000.

[6] F.Melgani and L.Bruzzone, " Classification of hyperspectral remote sensing images with support vector machines," IEEE Trans.Geosci.Remote Sens., vol.42,no.8,pp 1778-1790,2004

[7] Y.Tarbalka, M.Fauvel, J.Channusot, and J.Benediktsson, " SVM and MRF based method for accurate classification of hyperspectral images," IEEE Geo sci Remote sens.Lett., vol.3 no.7, pp.736-740,2010

[8] R.Battiti ," Using mutual information for selecting features in supervised neural net ., learning," IEEE Trans .Neural Nets., vol.5,no.4,jul.1994

[9] A.Gualtieri and R.F.Cromp, " Support Vector Machines for hyperspectral remote sensing classification,"Proc.SPIE,vol.3584,pp.221-232,Jan 1998

[10] B.Scholkopf,A.Smola,"Learning with Kernels-Support Vector Machines,Regularization,Optimization and Beyond,"MIT press series,2002

[11]. L. Kuncheva, "Combining Pattern Classifiers: Methods and Algorithms". Hoboken, NJ, USA: Wiley-Interscience, 2004.

[12]. J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 3, pp.226–239, 1998.

[13]. Andrey Bicalho santos, Arnaldo de Albuquerque Araújo, David Menotti, "Combining multiple classification methods for hyperspectral data interpretation" IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 6, no. 3, June 2013

K.Kavitha is currently working as Assistant Professor in ECE Department from Mepco Schlenk Engineering College, Sivakasi. Her research area is image classification and change detection.

G.Manochitra is a PG scholar from Mepco Schlenk Engineering College, Sivakasi. Her research area is image processing and signal processing